

**58th Stock Assessment Workshop/Stock Assessment  
Review Committee (SAW/SARC):**

---

**Independent reviewer's report on the Benchmark stock  
assessments for butterfish, tilefish, and northern shrimp**

---

**Stewart Frusher**

*Prepared for*

Center for Independent Experts

## Table of Contents

|   |    |
|---|----|
| Executive Summary .....   | 3  |
| Background .....  | 5  |
| Description of Stewart Frusher’s role in the Review Activities.....   | 5  |
| Review of the Terms of Reference.....   | 6  |
| Conclusions .....   | 23 |
| Recommendations in accordance with the TOR’s .....  | 24 |
| Critique of the NMFS review process, including suggestions for improvements of both process and products..... | 24 |
| Appendix 1: Bibliography of materials provided for review .....   | 26 |
| Appendix 2: A copy of this Statement of Work.....   | 28 |
| Appendix 3: Panel Membership or other pertinent information from the panel review meeting. ....               | 45 |

## **Executive Summary**

### **(i) Butterfish**

All TORs were accepted and the assessment results can be used for management advice. While the presented modified Age Structured Assessment program model (ASAP3) was not seen as the best configuration of the model for the current assessment, a refinement of the model that included removal of the habitat model and restricting the survey indices to only the Fall surveys for this assessment was appropriate. Outcomes from the refined model provided similar conclusions to the assessment outcomes in that the stock is neither overfished or that overfishing is occurring.

The habitat model was seen as being innovative and worthy of future research and should be considered in the future, however the outcomes from the model essentially showed no trends over the time series and thus the outcomes from the ASAP 3 model with and without the habitat model were very similar. Thus inclusion of the habitat model required a greater number of parameters to be estimated with no improvement in outcomes and the combined model (ASAP3 + Habitat) was therefore less parsimonious than the ASAP3 model only.

The opposite trends in the residuals from the Fall and Spring surveys together with the lack of correlation between the Spring and subsequent Fall survey (age 0 and age 1 fish) suggested that both surveys could not be included in the model. The Fall survey was considered the most appropriate, based on its coverage and that the catchability coefficient ( $q$ ) was externally estimated. Careful consideration of the fishery independent surveys and their potential to demonstrate trends in estimated 'de facto' abundances needs to be considered especially as butterfish is a pelagic species known to vertically migrate and that the surveys are demersal. Although too short to determine the patterns in the residuals, the preliminary data from the surveys using the *FSV Henry B. Bigelow* were encouraging.

Butterfish are mainly commercially distributed from Cape Hatteras to the Gulf of Maine. With increasing water temperatures, especially in northern regions, butterfish distribution and abundance would be expected to move northwards with declining abundances in southern regions of the distribution. However, there appears to be limited knowledge of inshore coastal/estuarine habitats and with the environmental changes being noted in the this region (e.g. Gulf of Maine), information from the State based surveys, especially as most of these are in shallower waters, should be an integral component in determining the distribution and abundance of butterfish for future assessments.

### **(ii) Golden Tilefish**

All TORs were accepted and the assessment results can be used for management advice. The presented ASAP model was appropriate with a minor adjustment as to how the recent recruits were being estimated for the projections. Outcomes from the model were consistent with the conclusion that overfishing is not occurring and that the stock has rebuilt.

TOR 3 was only partially accepted as it was felt that while a temperature profile had been developed,

there was limited discussion on how this could impact the stock or its potential use in the assessment. However, as no environmental variables are built into the model, a partial acceptance of this TOR does not impact on the management advice.

The assessment is primarily dependent on landings per unit effort data which relies on a relatively crude measure of effort – boat days. As there are few dedicated golden tilefish vessels it was considered that there is the opportunity to obtain more detail on effort such as soak time, number of hooks, length of longline etc. Similarly with few processors, improved standardization of market categories should be achievable. Refinement of the standardized catch data is recommended.

The fishery appears to be strongly recruit driven with frequent (every 5-6 years) relatively consistent positive recruitment periods. As there is a delay between recruitment to the stock and recruitment to the fishery (approximately 4 years), the last three estimated recruitments in the model are based on limited information and thus for projecting future catches it was considered that mean recruitment based on the historical estimates was more likely rather than higher or lower recruitment based on the past few years of positive or negative recruitment. Efforts to improve knowledge of recruitment, either through pre-recruit surveys or linkages to environmental drivers, are recommended.

A dome shaped selectivity curve was used to fit the model as few fish in the larger size categories would remain in the fishery with an asymptotic or logistic selectivity curve. There was anecdotal support for this from fisher representatives from the meeting and the spatial distribution of large and small fish. Further work to confirm the selectivity function is recommended.

### **(iii) Northern Shrimp**

The TORs associated with the performance of an appropriate model and the associated management advice were not met.

The University of Maine (UME) model is an appropriate model for difficult to age species and has been successful in crustacean fisheries globally. While the assessment team had done a great job in developing the model, it was felt that comprehensive testing of the model was still required. Of particular concern was the lack of fit to the survey indices and survey size composition which impacted the accuracy and precision of model outputs. Further development and validating of the model is recommended.

The Catch Survey Analysis (CSA) model had very poor fits to the survey indices. Of concern was the reliance of catch as a measure of abundance. As the fishery had experienced changes in fleets and fishing seasons, unstandardized catch information can reflect changes in effort as much as abundance. Adjusting the CVs in favour of survey indices and less reliance on catch made only marginal improvements and likelihoods were the same for alternative weightings indicating that management advice could change (e.g., from overfishing to not overfishing) based on the selected weights. As such, the model is not robust enough for management of the fishery. Utilising standardised catch information (e.g., catch per unit of effort) was recommended as a potential variable to be explored in future models.

The ASPIC model, which had been used in the past, is not appropriate as this model is unable to handle the strong year classes that are seen in this fishery.

The Gulf of Maine region is the southern extent of the commercial fishery for northern shrimp. As water temperature warms it would be expected that the distribution of the species moves poleward. Given the documented environmental changes occurring in the Gulf of Maine and the potential for recruitment peaks to be environmentally driven, it is recommended that further work on the biophysical relationships, including climate drivers and ecosystem interactions be explored.

## **Background**

The 58th SARC (Stock Assessment Review Committee) met in Woods Hole, MA from 27 January – 31 January, 2014 to review stock assessments for butterfish (*Peprilus triacanthus*), golden tilefish (*Lopholatilus chamaeleonticeps*), and northern shrimp (*Pandalus borealis*). The review committee was composed of Dr. Robert J. Latour (MAFMC SSC and Virginia Institute of Marine Science, Chair) and three scientists affiliated with the Center for Independent Experts: Dr. Cathy Dichmont (CSIRO), Dr. Stewart Frusher (University of Tasmania), and Dr. Ian Jonsen (Dalhousie University).

The SARC was assisted by the NEFSC SAW Chairman, Dr. James Weinberg, his staff, and Dr. Paul Rago (NEFSC). Supporting documentation for the butterfish assessment was prepared by the NEFSC Coastal/Pelagic Working Group, and presentations were made by Drs. Charles Adams, John Manderson, and Timothy Miller (NEFSC). Written material for the golden tilefish assessment was prepared by the NEFSC Demersal Working Group, and presentations were made by Dr. Phil Nitschke. Technical documents for the northern shrimp assessment were prepared by the ASMFC Northern Shrimp Technical Committee, and presentations were made by Ms. Kelly Whitmore and Drs. Anne Richards and Katie Drew.

The previous benchmark stock assessment for butterfish was in 2009 (SARC 49, NEFSC 2010), tilefish was also in 2009 (SARC 48) and northern shrimp in 2007 (SARC 2007). Since these last assessments considerable research advancements have been made in each assessment including the incorporation of new data and refinement of modelling approaches. The assessments conducted by the CPWG, DWG, and NSTC were very thorough, and it was apparent that each group devoted significant time and effort to data analysis, model fitting, evaluation of uncertainty, and report preparation.

## **Description of Stewart Frusher's Role in the Review Activities**

My role in the 58th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC) was to read all the documentation provided (as listed in Appendix 1) prior to attending the review panel from the 27<sup>th</sup> to 31<sup>st</sup> January 2014 in the Stephen H. Clark Conference Room – Northeast Fisheries Science Center, Woods Hole, Massachusetts. During the meeting I heard presentations on the assessments for each of the species and discussed, along with the other panel members, the

suitability of the models to provide management advice. On the final day (Friday 31<sup>st</sup> January) I participated in a closed session with the panel chair and the other two reviewers to begin drafting the agreed panel summary report. After leaving Woods Hole, I reviewed the agreed panel summary report and prepared my individual report.

## **Review of the Terms of Reference (TOR)**

### **(i) Butterfish**

#### **1. Characterize the commercial catch including landings, effort and discards by gear type.**

##### **Describe the magnitude of uncertainty in these sources of data.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The difficulty with this fishery is that there are multiple sources of catch data from a range of vessels (national and foreign) and gear. Additionally, the fish has been and still is part of a mixed fishery (no specific targeting) and discarded fish have been a large part of the mortality profile. While the fishery has a long history (1887 – current), most of the earlier data have considerable uncertainties making them problematic to use. Truncating the series to periods where there was improved certainty in the data being used in the model is acceptable and the choice of 1989 as the starting point has considerable merit as outlined in the report and I fully support this choice. However, having chosen a cut off, it should not result in future assessment teams ignoring the earlier data as there is considerable semi-quantitative information in, for example, the magnitude of previous catches.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The data and associated uncertainties are sufficient to provide scientifically credible basis for developing fishery management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. During the review it was mentioned that butterfish were occurring (occasionally?) further north in Newfoundland. As the Gulf of Maine region warms, it might be expected that the distribution of this species moves further poleward. As such, it would be beneficial to try to encompass catch information from northern regions as an indicator that the species distribution could be moving further north. The temperature-habitat model results indicated a very flat trajectory despite the region showing considerable warming. If this is correct then it would be expected that the species would move with its temperature habitat envelope.

#### **2. Characterize the survey data that are being used in the assessment. Describe the magnitude of uncertainty in these sources of data.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. Determining the status of this resource is very dependent on fishery independent data due to the inconsistency in the availability of fishery data. Despite there being a

large number of surveys available, it was disappointing that many of these were dismissed as they didn't provide large spatial coverage and different methods were used in different states. However, standardisation of these surveys would provide additional data that could be used to support the main NEFSC data. This is of particular importance when the inshore and offshore surveys are showing opposite patterns in the residuals and there was a negative correlation between the fall survey (age 0 fish) and the following spring survey (age 1 fish). Thus it became awkward to determine which survey was the best to use. The spring survey tended to follow cohorts slightly better but provide relatively limited spatial coverage compared to the fall survey and the numbers of fish caught during the spring survey was also substantially lower, which had led to higher CVs.

- b. The State based surveys are also expected to provide insights into the close inshore regions where the new research vessel (FSV *Henry B. Bigelow*) has limited capacity to operate compared to the previous vessel (FSV *Albatross IV*). The inshore and estuarine regions appear to be an important component of the life history of butterfish.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The information on the trends and uncertainty provided by the CPWG was acceptable and is essential for developing management advice. Further detail regarding the use of the survey data is addressed in TOR 5.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. Despite the CVs of the survey data improving with the use of the new vessel, there is concern that butterfish is a pelagic fish species that undertakes daily vertical migrations and thus a bottom trawling vessel is likely to generate bias in the data being collected. The position of vertically migrating species is often determined by the time of day (e.g. changing light conditions at dawn and dusk often trigger migrations) and the weather conditions (e.g. fish may remain closer to the surface or remain in surface waters for longer when the weather is darker due to cloud cover; fish may aggregate into tighter schools during periods when the day is "brighter"). A GAMS using variables such as time since dawn, time to dusk, cloud cover should be explored to determine if there are physical variables that can be used to standardise the data further.

### **3. Characterize oceanographic and habitat data as it pertains to butterfish distribution and availability. If possible, integrate the results into the stock assessment (TOR-5).**

- This TOR was met.
- Strengths/Weaknesses.
  - a. Pelagic species are directly (e.g. physiology) and indirectly (e.g. prey availability) dependent on the physical characteristics of the water mass for their productivity (cf habitat structure for demersal fish species). There are a number of variables associated with the pelagic environment that can affect the distribution and abundance of pelagic species – temperature, nutrient supply, productivity, currents, light penetration, stratification etc. Ideally all these data sources would be

incorporated into a habitat preference model, although the availability and quality of data needs to be determined. However, if one was to pick a single variable then temperature would be the obvious choice and knowledge of the impact of temperature on the species distribution, abundance and availability is important given the rising temperatures being experienced in this region.

- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. Although the habitat model was not incorporated into the final model that was accepted (see TOR 5). This was not a reflection on the quality of the model, rather the model outputs currently provided limited contrast (i.e. were relatively flat across years) and thus the additional parameters required to estimate the habitat model resulted in the combined model being less parsimonious than the model without the habitat model.
  - b. An interpretation of the lack of contrast in the habitat model would be that the distribution of butterfish is tracking the species temperature range exceptionally well. This is more likely to be the case for a pelagic species although the substantial changes in abundance (recruitment) together with the substantial changes in the temperature of the region do suggest that further investigation of the model is required. Unfortunately there was insufficient time to do this in the time allocated.
  - c. The method is innovative and it would be advantageous to see it explored more. Developing it for a demersal species where habitat structure is known to contribute to distribution and abundance may help demonstrate its potential usefulness.
  - d. If the model does track availability correctly, then it would be a great asset for assessment models and enhance any model's ability to incorporate climate change signals. Given that the Gulf of Maine is considered a climate change hotspot, further work in this area is encouraged.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. While temperature is a key variable to use, pelagic fish distribution, reproductive quality and abundance are also often linked to productivity. Richards et al (2013 – see Northern Shrimp documentation) demonstrated that there have been changes in productivity in the Gulf of Maine due to increased runoff associated with changing rainfall patterns. Productivity, as measured by chlorophyll a concentrations from satellite images, is a potential additional parameter to consider in defining pelagic habitats, especially regions of abundance. Given that the GAM failed to converge during periods of high survey abundances and that temperature only explained 30% of the deviance, there is likely to be “something else” contributing to distribution and abundance and productivity (assuming equivalent to prey availability) should be explored.

**4. Evaluate consumptive removals of butterfish by its predators. If possible, integrate results into the stock assessment (TOR-5).**

- This TOR was partially met.
- Strengths/Weaknesses.



- a. Small pelagic fish species often play important roles in marine ecosystems as they are an important connection between primary productivity and higher trophic level species.
- b. NEFSC is fortunate to have a time-series database of stomach contents of higher trophic level predators of butterfish and the concept of using this information to adjust natural mortality rates is excellent.
- c. Stomach content analysis by visual examination of stomach contents does have acknowledged biases and a recent study has shown that there can be considerable under-estimation of key prey species (Barnett et al., 2010).
- d. With the move towards ecosystem based fisheries management (EBFM) over the past two decades, there has been a concomitant development of large and sophisticated ecosystem models (e.g. EwE, Atlantis). Although the functional groups in these models may not separate butterfish from other small pelagics, it is recommended that the ecosystem models be explored to determine if there is similarity in outcomes between approaches and to gain an understanding of the role of the predators that were not examined (sharks, seabirds, other predatory fish species, etc.).
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. It was considered that further work on predator removals would be required before they could be used to adjust natural mortality estimates in the assessment model.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. I am aware that there is considerable research, both underway and planned, to address the changing environmental conditions in the Gulf of Maine. Linkages and co-ordination across all programs and projects would be beneficial in understanding the impacts of the environment on the dynamics of this fishery.

**5. Use assessment models to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a comparison with previous assessment results and previous projections.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The ASAP model is an appropriate model for this species and has been extensively used and tested.
  - b. Several indices are available and the initially the model was fitted to six indices: fall inshore and offshore, spring inshore and offshore and the Bigelow offshore and inshore. The latter was too short a time series to be able to comment on although the CVs do look promising for its use in future assessments as the time series increases.
  - c. A major issue with the indices was the distinctive patterns in the residuals which were in contrast between the spring and fall surveys. Thus fitting to both surveys would “average” and produce no trend.
  - d. A correlation analysis undertaken on request from the review panel of age 0 in spring and age 1 in the following fall was negative. Neither survey tended to have a

strong relationship (correlation) between years which was surprising given that the species is short lived and there has been a considerable fluctuation in recruitment. The spring survey tended to following cohorts better than the fall survey. However, the fall survey is undertaken when butterfish are more broadly distributed across the survey area; catches were higher and provided greater contrast in the data.

- e. Patterns in the residuals were also apparent in the age composition data indicating that this was also not a good fit. This pattern was both temporal (e.g. negative in the late 1990s and then positive in all subsequent years for the NEFSC-spring-offshore survey) and between age classes (e.g. negative for age 1 and positive for age 2). The fall offshore and inshore surveys had a better fit to the age composition –possibly reflecting the greater sample size.
- f. Currently, only the fall survey estimated catchability (“q”) externally to the assessment model which was an innovative development and appropriate for a short lived species where “q” is likely to vary due to the survey type and seasonal variation.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. Because of the conflicting results from the different surveys only one survey should be used. While the spring survey tended to follow cohorts better, the fall survey was considered a better option at this stage as the age composition residuals and CVs were better fits, the numbers caught per tow were greater and provided the ability for the model to estimate M. Thus the use of ASAP3 with the fall only survey was the most appropriate. It was also considered that the habitat index offered no further advantage (at this stage) and model runs with and without the habitat index showed similar results. While this index remains flat (i.e. shows no trend or variation that could explain variation in other variables in the model), there is no reason to incorporate it into the model, although see TOR3.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. Given the concerns over the catch and discards data, the survey data and the age composition data are key inputs for this assessment. As a pelagic species that schools and vertically migrates, there is concern that a bottom trawl does not capture the behavioural dynamics of the species and this may be contributing to the lack of consistency across seasons and between years (see TOR 2). Further efforts at understanding how well these benthic surveys reflect butterfish abundance is recommended.

**6. State the existing stock status definitions for “overfished” and “overfishing”. Given that the stock status is currently unknown, update or redefine biological reference points (BRPs; point estimates for  $B_{MSY}$ ,  $B_{THRESHOLD}$ ,  $F_{MSY}$  and  $MSY$ , or their proxies) and provide estimates of their uncertainty. Consider effects of environmental factors on stability of reference points and implications for stock status.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The BRPs are based on standard BRP surrogates which were both acceptable and precautionary.

- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The BRPs were determined for management advice although greater efforts in evaluating model sensitivities would be beneficial and it is recommended for future assessments.
  - b. As the final accepted model did not have the habitat index, environmental variability was not assessed.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. The environmental factors were addressed through the habitat index approach. While this was an innovative and sophisticated approach and worthwhile of further development, there are other indices that could also be considered including ENSO, NAO etc. Given the changing observed environmental conditions in the Gulf of Maine region, understanding linkages to environmental variables is important and a series of indices should be considered.

**7. Evaluate stock status with respect to a newly proposed model and with respect to “new” BRPs and their estimates (from TOR-6). Evaluate whether the stock is rebuilt.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. Similar to TOR 6
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The available information and model outputs provide sufficient indications that the stock is neither overfished or that overfishing is occurring.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. No further clarification required for this TOR.

**8. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).**

- a. **Provide numerical annual projections (2 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment). Comment on which projections seem most realistic.**
- b. **Describe this stock’s vulnerability (see “Appendix to the SAW TORs”) to becoming overfished, and how this could affect the choice of ABC.**
- This TOR was met.
- Strengths/Weaknesses.
  - a. Using model estimated hindcast recruitment indices is standard although consideration needs to be given to any patterns in these indices. For example, if all the latter years were low (possibly reflecting an environmental impact), then projecting forward using a combination of high and low estimates over a longer time series could artificially inflate biomass projections. However, there appeared to be no such trend in the butterfish data and use of the entire time series is appropriate.

- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The results provide a scientifically credible basis for developing fishery management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. No further clarification required for this TOR.

**9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The model for this species is dependent on fishery independent indices as the fishery (including discards) is currently relatively minor. As such, given the concerns over the opposite patterns in the residuals for the different surveys and the weak ability to track cohorts, future emphasis should be on understanding these indices. Thus the first recommendation to examine efficiency and catchability of the gear is essential.
  - b. Given the growing information about the changing environmental conditions in this region, it is recommended that further efforts in understanding the relationships between the environment and butterfish population dynamics be explored. As a small pelagic species, understanding the role of butterfish in the ecosystem is important, including the potential impacts that removal of butterfish biomass may have on other commercial or keystone ecosystem species.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
 

The four recommendations by the CPWG are all appropriate. The two recommendations regarding the survey indices (1) and the environmental covariates (3) would improve future assessments and thus management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. Given the importance of the fishery independent indices as discussed above, it is important that either consistent indices are used or that the most appropriate indices be used. Calculating “q” outside the model as done for the fall survey is appropriate and this needs to be done for the spring survey so that the ASAP3 model with external estimated “q” for spring only surveys can be compared to the fall only surveys.
  - b. Although limited in spatial coverage and using different gear types, the state surveys should be analysed either through a statistical approach that combines all indices or as a means to compare trends and patterns in an attempt to improve the fishery independent indices or help in validating which of the fall and spring indices is the most appropriate to use.
  - c. As previously mentioned, greater emphasis should be placed on working with the ecosystem researchers.

## B. Golden Tilefish

### 1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the magnitude of uncertainty in these sources of data.

- This TOR was met.
- Strengths/Weaknesses.
  - a. The assessment team had done a good job in bringing together all the catch data.
  - b. Restricting the time series to the period when the directed fishery started is appropriate given the uncertainties in the earlier data.
  - c. Currently the limited number of fishers and processors provide a unique opportunity to obtain precise spatial and temporal data. As indicated by the DWG, there is inconsistency in the reporting of market categories and efforts to rectify this should be relatively simple given the number of processors.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The catch data from the early 1970's when directed fishing began is adequate to support the assessment and provide management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. No further clarification required for this TOR.

### 2. Characterize commercial LPUE as a measure of relative abundance. Consider the utility of recreational data for this purpose. Characterize the uncertainty and any bias in these sources of data.

- This TOR was met.
- Strengths/Weaknesses.
  - a. Standardised forms of commercial catch provide the best index of abundance and underpin many fishery assessments. While catch is normally easily determined from logbooks, landings and/or processor statements, the standardisation term (effort) is more difficult. Trip length, as used in this assessment, is probably the least desirable as it requires each fishing day to be consistent. Variability can occur when different numbers or sizes of hooks are used, different lengths of longline, different soak times (can vary considerably as hours of daylight varies throughout year) etc.
  - b. Standardised forms of commercial landings also fail to reflect abundances when catch is restrained (e.g. by quota or markets). Generally catch rates are inflated and biomass estimates are over inflated. This is because fishes target dollars per trip rather than catch per trip. This fishery has a quota and there was anecdotal reporting from the fishing industry representative at the meeting that fishing was being undertaken to prevent flooding of the market and to maintain a higher price.
  - c. Despite the concerns above, the DWG have undertaken a detailed analysis of CPUE comparing various estimates from earlier studies and the newer VTR information. The development of a GLM that includes year and vessel effects is good and the consistency between individual vessel CPUE and between GLM and nominal CPUE is sufficiently encouraging to suggest that CPUE is adequate.

- d. Given the environmental changes that are being recorded in this region, it would be advantageous to consider larger scale physical drivers such as temperature, currents, productivity changes etc. This is particularly important for this species where frequent recruitment peaks are observed which have substantial impacts on CPUE data.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The commercial LPUE is adequate for use in the assessment to determine management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. There is a very good opportunity to work with industry in this fishery to improve both the effort information and also environmental information. For example, depth and temperature loggers are relatively cheap and could be attached to the longline. GPS locations of set fishing gear with respect to bottom profiles could be obtained provided “trust” can be built with the industry so that “specific” locations are not presented in public documents that identify a specific fishers favourite fishing locations.

**3. For the depth zone occupied by tilefish, examine the relationship between bottom temperature, tilefish distribution and thermal tolerance.**

- This TOR was partially met.
- Strengths/Weaknesses.
  - a. A temperature profile for the species was developed and provides a starting point for further ecosystem studies. However, while the report acknowledges that cold water may have resulted in a “massive” golden tilefish die off in the early 1880s and that this may have been attributed to a low NAO, no attempts were made to link abundance with any physical oceanographic features. Additionally, there are consistent episodic recruitment patterns which are also likely to be environmentally driven. Given the emphasis being placed on this region due to its changing environment, there would be benefit in linking ecological and stock assessment models.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. No temperature profiles or effects are incorporated into the model and thus there is no ecosystem assessment or links with environmental and climate drivers that can be used to provide management advice at this time.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. In general this area was very poorly understood which provided the assessment team with limited information to incorporate into the assessment. There was anecdotal evidence that there is a spatial component to the tilefish distribution with smaller fish being in shallower water. Given the narrow thermal range, catch rates could be substantially altered by changes in bottom water temperature – either by allowing the stock to disperse or by concentrating it in specific regions.

**4. Use assessment models to estimate annual fishing mortality and stock size for the time series, and estimate their uncertainty. Include a historical retrospective to allow a comparison with previous assessment results.**

- This TOR was met, although the last part of the TOR (comparison with previous results) could not be undertaken as this was the first time the ASAP approach had been used.
- Strengths/Weaknesses.
  - a. Model comparison was a strength of this assessment as it highlighted where model improvements could be made as well as where there were consistencies across models. Greater confidence was assigned to the ASAP model as improvements could be easily documented in comparison to the SCALE and ASPIC models.
  - b. The golden tilefish fishery has several unique properties which were better suited by the ASAP model. These included frequent and relatively periodic periods of recruitment, a long life history and a relatively narrow harvest age. The later was addressed through a dome-shaped selectivity curve. Although not common in assessments there was evidence from the fishers that older age classes were not targeted as they received a lower price, were in more difficult habitat to fish and in certain regions they clashed with other fisheries (i.e. gear conflicts).
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The ASAP model provided reasonable fits to the data and was appropriate for providing fishery management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. N/A

**5. State the existing stock status definitions for “overfished” and “overfishing”. Then update or redefine biological reference points (BRPs; point estimates for  $B_{MSY}$ ,  $B_{THRESHOLD}$ ,  $F_{MSY}$  and  $MSY$  or for their proxies) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the “new” (i.e., updated, redefined, or alternative) BRPs.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. Given the positive increase in SSB under a constant quota stretching over a decade, there is merit in using the average  $F$  for this period as a proxy for  $F_{MSY}$ . This was then used to estimate  $SSB_{MSY}$  and  $MSY$ .
  - b. Over the 10 year period of the constant quota the stock has rebuilt. During this period there has also been several high and low recruitment periods indicating that the approach for estimating  $F$  is relatively robust.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The new BRPs are appropriate for providing management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. No further clarification required for this TOR.

6. **Evaluate stock status with respect to the existing ASPIC model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt.**
  - a. **When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.**
  - b. **Then use the newly proposed model and evaluate stock status with respect to “new” BRPs and their estimates (from TOR-4).**
  - This TOR was met.
  - Strengths/Weaknesses.
    - a. Due to the lack of fit of the ASPIC model to the observed data and the poor ability of the ASPIC model to handle the variable recruitment that is found in this fishery, there was little support for the BPR’s estimated previously using the ASPIC model. The ASAP model is structurally different to the ASPIC model and has a substantially better fit to the observed data including the recruitment variations.
    - b. The conclusion made by the DWG that the stock is rebuilt is considered appropriate.
  - Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. The BRPs and stock status determined from the ASAP model are all scientifically credible and can be used for management advice.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. No further clarification required for this TOR.
7. **Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).**
  - a. **Provide numerical annual projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for F, and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).**
  - b. **Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.**
  - c. **Describe this stock’s vulnerability (see “Appendix to the SAW TORs”) to becoming overfished, and how this could affect the choice of ABC.**
  - This TOR was met.
  - Strengths/Weaknesses.
    - a. The fishery tends to have fairly consistent recruitment peaks with positive and negative deviations lasting from 2-4 years. Residuals from the last 6 years have all been negative which is longer than the previous patterns. However, the last 3 years are based on virtually no data as golden tilefish do not recruit to the fishery until at least age 4. Given the trends observed in recruitment it is likely that there had been positive recruitment to the fishery but this was yet to appear in the data. The fishers at the meeting confirmed that they were seeing smaller fish and one fisher, who was



also a processor, provided his latest market categories which indicated smaller fish were entering the fishery. The latest research samples obtained from the fishery were also showing signs of smaller fish, but these were not available at the time of the assessment. While anecdotal statements such as these cannot be used in the assessment process, the DWG undertook an empirical adjustment of the last three years of the recruitment estimates to equal the mean value during the assessment period. While this resulted in higher estimates of biomass, this outcome appeared to be a better expected fit to the rebuilding trends seen in the fishery since the quota was introduced. Furthermore, the pre-adjusted figures resulted in overfishing of the stock in the first few years of projections which was at odds with the dynamics of the stock under the quota regime. Thus the new estimates of biomass are consistent with the other indices from the fishery.

- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The projections with the adjusted recruitment estimates appear suitable for management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. In accessing a stock of this nature where there are no fishery independent data and concerns over the CPUE data then “multiple lines of evidence” are often considered. Importantly for this stock there has been a range of good and poor recruitments over the quota period (since 2002) and biomass has improved over this period. Closer examination of the recruitment deviations (Fig. B87) show that there has been more negative residuals due to 2 poor recruitment periods than positive recruitments (one in 2006). As such, favourable or unusual recruitment would not be biasing estimates. Similarly, if the recruitment cycles continue into the future similar to the past, then the fishery should be expecting a positive recruitment cycle and there is anecdotal evidence for this as mentioned above. Thus all information – both anecdotal, observed and estimated - all indicate that the stock’s vulnerability is low and, importantly, estimates of *SSB* and *F* are conservative.

**8. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The first of the research recommendations provided by the DWG– development of an industry based survey similar to a fishery independent survey and increased maturity sampling, is strongly support. However, given the dependence on recruits to the fishery (i.e. CPUE is strongly tied to recruitment peaks), consideration should be given to trying to obtain information on pre-recruits to the fishery. This could involve sampling shallower or with different gear (smaller hooks?).
  - b. The second research recommendation: increase maturity sampling seems odd given that there has been a study undertaken that has provided maturity information for the golden tilefish. However, I would suggest that there is a need to improve maturity information but not from a reproductive biology perspective, but from a

functional perspective. Although very speculative, the consistency in the frequency of recruitment to this fishery during periods when the region is experiencing substantial environmental change may suggest that there is a functional reproductive period where by there is significant spawning success every 5-6 years.

- c. In conjunction with (b), there is a need to investigate large scale climate and environmental drivers on abundance and recruitment. Given the strong relationship between recruitment and CPUE, any advancement in the development of pre-recruit indices either through sampling (see 'a' above) or correlations with environmental variables would be an early warning signal for this fishery.
  - d. The aging information has been a significant step and is a key input into the ASAP model. Validation of the ages needs to be undertaken and continued ageing into the future.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. All the research recommendations would provide increased certainty in the assessment outputs that would then flow into management of the fishery.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. The small number of dedicated fishers (4) and processors provide a great opportunity to work with industry to obtain precise and robust data. In particular, there is the opportunity to gain improved information on the LPUE such as soak time, number of hooks, environmental conditions, spatial coverage (GPS and depth) and the market categories, so that there is consistency in the data being provided for the assessment.

### C. Northern shrimp

1. **Present the Gulf of Maine northern shrimp landings, discards, effort, and fishery-independent data used in the assessment. Characterize the precision and accuracy of the data and justify inclusion or elimination of data sources.**
- This TOR was partially met. Disagree with the exclusion of the CPUE data and use of catch data. Otherwise all other components of the TOR are accepted.
  - Strengths/Weaknesses.
    - a. The assessment team had provided estimates of commercial catches from 1960 to 2013. During this period the catches were highly variable. Such variability can be due to recruitment variability or effort changes or a combination of both. Given both the change in the number of vessels operating in the fishery and the long list of management changes imposed that impact catch (seasonal closures; Table C4.1) and the effects of market (prices), catch is likely to be an unreliable measure of abundance in the fishery. With such variability in the catch data, standardisation of catch as CPUE is normally used and CPUE is normally seen as a more accurate reflection of changes in abundance. A correlation between landings and CPUE (Maine pounds per hour towing; Table C5.9) was weak ( $r=0.24$ ) suggesting that either catch or CPUE were not reliable indices.

- b. Fishery independent data also provide measures of abundance. Only the ASMFC and the NEFSC Fall summer indices were used and while the justification for this choice was reasonable (spatial and temporal issues), the other indices would contain information on changes in relative abundances and recruitment, and effort should be made to look at ways of standardising these for inclusion in future assessments or using them to interpret environmental impacts on the dynamics of the stock. Given the tight confidence limits around the northern shrimp survey indices and the concerns about the catch and effort data (see 'a' above) future assessments are likely to be reliant on these data (but see 'c' below).
- c. Given the concerns over the catch and effort data, it is reasonable to determine how well either catch or CPUE match the ASMFC and NEFSC indices. The table below shows the correlation coefficients for these comparisons. As expected the CPUE for the inshore, offshore and combined are highly correlated. As stated above (see 'a') there is limited correlation between the landings and CPUE and no correlation between landings and any of the fishery independent indices. In contrast the CPUE has moderate correlations with all survey indices indicating that CPUE is likely to be a much better proxy for abundance than catch.

| Index                               | CPUE – inshore<br>(Table C5.9) | CPUE –offshore<br>(Table C5.9) | Landings<br>(Table C5.1) |
|-------------------------------------|--------------------------------|--------------------------------|--------------------------|
| CPUE - inshore                      | 0.86                           | 0.98                           |                          |
| CPUE - offshore                     |                                | 0.92                           |                          |
| Landings                            |                                | 0.24                           |                          |
| Shrimp index (>22)<br>(Table C13)   |                                | 0.49                           | 0.01                     |
| Shrimp index (total)<br>(Table C13) |                                | 0.50                           | -0.01                    |
| NEFSC Fall Survey<br>(Table C5.15)  |                                | 0.60                           | -0.06                    |

- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. With the exception of the exclusion of the CPUE data and the use of catch as a de facto abundance measure, the other data are suitable for the assessment.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. No further clarification required for this TOR.
2. **Estimate population parameters (fishing mortality, biomass, and abundance) using assessment models. Evaluate model performance and stability through sensitivity analyses and retrospective analysis, including alternative natural mortality (M) scenarios. Include consideration of environmental effects where possible. Discuss the effects of data strengths and weaknesses on model results and performance.**
- The TOR was not met.
  - Strengths/Weaknesses.
    - a. The UME model is an appropriate approach for crustacean fisheries and has been used extensively in lobster fisheries. However, before use, extensive simulation

testing needs to be undertaken and it appeared that while considerable work had gone into developing the model for northern shrimp, these tests were still underway. Further development of the model is encouraged including peer-review before its use to provide management advice. Of particular concern was the lack of fit to the survey indices and the size composition data. Given that these are considered the most robust data available for this assessment (see TOR 1), it is unlikely that the model, in its present form would provide scientifically credible management advice.

- b. The CSA model also had difficulty fitting to the data with strong cyclic trends in the survey index residuals. A large problem with the data is the very large index for 2006 which completely disappears in 2007. This peak was not reflected in the 2006 catch and a representative from the floor indicated that initial high catch rates resulted in flooding of the market and a subsequent drop in price that had some fishers not going to sea. This anecdotal information also supports the preferred use of CPUE data compared to catch data. To try to improve the fit of the model the weights of the catch data were relaxed as were survey indices. However, although the model was run with substantially different weightings of catch and survey indices, there was limited improvement in the fit to the data with the likelihoods being equal for these runs. However, *F* estimates varied enormously, indicating that vastly different management advice could be obtained depending on the choice of weights – all of which were equally likely. As such, the current northern shrimp CSA model was not robust enough to provide management advice.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. As discussed above, neither of the models was sufficiently robust to provide management advice.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. It is worth reiterating that it is not uncommon for a model to be performing adequately with a shorter time series of data and then to perform poorer with additional data, as was the case in this assessment. As such, the rejection of the current CSA model does not reflect on the SARC 45 assessment.
  - b. Unfortunately, with the issues identified in the model and the re-runs requested there was insufficient time to discuss environmental effects. However, it was noted that temperature had been correlated with northern shrimp recruitment and that the Gulf of Maine is undergoing considerable environmental fluctuations with evidence of significant warming. As the northern shrimp is a cold water species at the southern end of its distribution, it is not unreasonable to expect greater variability in recruitment and survival, as well as a general decline in abundance with time. As such, further work on environmental variables and northern shrimp stock dynamics is encouraged. Future modelling should either incorporate an environmental parameter that scales recruitment or natural mortality (*M*) or allow for either recruitment or *M* to be estimated in a separate environmental model prior to inclusion in the assessment model.

3. **Update or redefine biological reference points (BRPs; point estimates or proxies for  $B_{MSY}$ ,  $SSB_{MSY}$ ,  $F_{MSY}$ ,  $MSY$ ). Evaluate stock status based on BRPs.**
  - This TOR was not met.
  - Strengths/Weaknesses.
    - a. As no assessment model adequately addressed the data, no BRPs could be evaluated.
    - b. With no model to guide management advice, current management advice needs to use a “weight of evidence” approach that could include the survey time series, the CPUE and, to a lesser extent, landings.
    - c. Given the current state of the fishery that includes low survey indices, low CPUE and low landings and the planned 2014 fishery closure, there is a need to re-evaluate the adequacy of the current northern shrimp BRPs.
  - Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. BRPs were unavailable to provide management advice.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. Further concern over the performance of the model was the very low  $F$  value. This is unusual in a fishery where the industry struggle to catch the quota and the catch rate was also low.
4. **Characterize uncertainty of model estimates of fishing mortality, biomass and recruitment, and biological reference points.**
  - This TOR was not met.
  - Strengths/Weaknesses.
    - a. This TOR can only be addressed if a model (TOR 2) was accepted.
  - Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. N/A.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. N/A.
5. **Review the methods used to calculate the annual target catch and characterize uncertainty of target catch estimates.**
  - This TOR was not met.
  - Strengths/Weaknesses
    - a. Without an accepted model (TOR 2), this TOR cannot be addressed.
  - Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
    - a. N/A.
  - Elaborate on any points raised in the SARC Summary Report for further clarification.
    - a. N/A.

**6. Develop detailed short and long-term prioritized lists of recommendations for future research, data collection, and assessment methodology. Highlight improvements to be made before the next benchmark assessment.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The NSTC has developed a thorough list of research recommendations under 5 broad areas and has prioritised the research in each of these areas. In all there are 32 priorities of which 16 are high, 15 moderate and 1 low. Despite these priorities, of the six improvements that are to be made before the next benchmark assessment, only 50% of these are high priorities.
  - b. As these recommendations were made before this review, there was no indication that the models would be rejected. Thus a key high priority recommendation is to develop an acceptable model for assessing the stock. Both the CSA (short-term) and the UME (long-term) should be progressed and the assessment team would be encouraged to explore what effort indices might be available to standardise catch. In developing future models the NSTC may wish to incorporate bio-economic and economic management models and the move towards *MEY* reference points.
  - c. Given that the species is short lived, has variable recruitment, is at the southern limit of its commercial distribution and has known temperature preferences, there is a need to understand the impact of the environmental and climate drivers on future abundance and distribution of the stock in the Gulf of Maine region.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. The prioritised recommendations would, once complete, provide data that would be beneficial in the future management of the stock.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. No further clarification required for this TOR.

**7. Based on the biology of species, and potential scientific advances, comment on the appropriate timing of the next benchmark assessment and intermediate updates.**

- This TOR was met.
- Strengths/Weaknesses.
  - a. The next benchmark assessment needs to be undertaken once a model, acceptable to the NSTC, has been developed.
  - b. As this model could take time to be developed, it is recommended that the NSTC consider basing management advice on a “weight of evidence” from the observed survey indices and CPUE data once the fishery recommences.
- Was the TOR completed successfully (i.e. does the work provide a scientifically credible basis for developing fishery management advice)?
  - a. N/A.
- Elaborate on any points raised in the SARC Summary Report for further clarification.
  - a. Given that there will be no fishery in 2014, it is essential that the fishery independent surveys are continued as these will provide the basis for management decisions on the future of the fishery.

## References

- Barnett A, Redd KS, Frusher SD, Stevens JD, Semmens JM (2010) Non-lethal method to obtain stomach samples from a large marine predator and the use of DNA analysis to improve dietary information. *Journal of Experimental Marine Biology and Ecology* 393: 188-192.
- Richards RA, Fogarty MJ, Mountain DG, Taylor MH. 2012. Climate change and northern shrimp recruitment variability in the Gulf of Maine. *Marine Ecology Progress Series* 464:167-178. 12 p.

## Conclusions

- The butterfish stock is currently lightly fished and not overfished and there are no signs of overfishing. Until there is a direct target fishery or substantial increases in discards, there is unlikely to be a change in the status of this fishery. There is concern over the representativeness of the fishery independent surveys, which are the main indices for assessing the fishery. Further research on these is required, especially as the fishery independent surveys are demersal surveys and the species is a small pelagic. Small pelagics are often important ecosystem species and understanding the role of this species in the broader ecosystem is encouraged.
- The golden tilefish stock has rebuilt and is not overfished and there are no signs of overfishing. The main input into the assessment is fisheries dependent information and there is concern over the basic nature of this index as a measure of abundance. Landings per day can be affected by different fishing practices independent of changes in abundance and it is recommended that further effort be directed at improving this metric. The fishery is characterised by frequent and relatively consistent recruitment peaks and attempts to obtain pre-recruit information either through juvenile surveys or correlations with environmental variables is encouraged.
- The northern fish stock could not be assessed but all indicators suggest that the stock is not in a healthy state and the current moratorium is supported. Further work on both the CSA and UME models are recommended. Given the changes in effort associated with changes in the number of vessels operating and the fishing season, it is unlikely that catch is a reliable measure of abundance. Efforts to standardise catch are recommended as a model input. This fishery operates on the southern region of the species distribution and with the substantial changes being observed in the Gulf of Maine, further research linking productivity of the species to climate and environmental drivers is recommended.

## Recommendations in accordance with the TOR's

| Species         | TOR | Recommendation   |
|-----------------|-----|------------------|
| Butterfish      | 1   | Accept           |
|                 | 2   | Accept           |
|                 | 3   | Accept           |
|                 | 4   | Partially accept |
|                 | 5   | Accept           |
|                 | 6   | Accept           |
|                 | 7   | Accept           |
|                 | 8   | Accept           |
|                 | 9   | Accept           |
| Golden Tilefish | 1   | Accept           |
|                 | 2   | Accept           |
|                 | 3   | Partially accept |
|                 | 4   | Accept           |
|                 | 5   | Accept           |
|                 | 6   | Accept           |
|                 | 7   | Accept           |
|                 | 8   | Accept           |
| Northern Shrimp | 1   | Partially accept |
|                 | 2   | Reject           |
|                 | 3   | Reject           |
|                 | 4   | Reject           |
|                 | 5   | Reject           |
|                 | 6   | Accept           |
|                 | 7   | Accept           |

## Critique of the NMFS review process, including suggestions for improvements of both process and products.

All three assessments were considerable in a relatively short period of time. Overall there were 24 TORs to be considered. This was inadequate as it resulted in:

- (i) The presenters for each assessment being rushed. There was a considerable volume of material being presented and it was hard to delve into detail as we were constantly aware of the timeslots for the agenda.
- (ii) With the problems associated with the assessment models, re-runs and critique of the models filled most sessions. There was limited time to discuss the biology of the species, the ecosystem in which it operates or the interpretation of the model results in light of the biology and ecology of the species. This was a major issue given that the Gulf of Maine region is experiencing major changes which are likely to impact all three species.



- (iii) Although there were options to have closed sessions so that the review panel could deliberate over points of discussion, the volume of material presented didn't allow for such breaks to occur as the reviewers were aware of the limited time available.
- (iv) By Thursday we were still having presentations by the three assessment teams which left little time to undertake the panel consensus review report which was partially completed (and rushed) on the Friday.

Overall, the assessment teams did an excellent job in preparing and presenting the materials. However, the northern shrimp assessment appears to be a combination of at least two different sets of material which had been rushed together as figures were incorrectly referenced. The report also arrived a week prior to the meeting. There was a lot of detail in the report and it is easy to overlook specifics, but the residual plots in figure C6.46 for the NEFSC Fall and ASMFC Summer surveys appear to be calculated differently (O-E; E-O). This was further reflected on the day when the developer of the UME model appeared unaware of the poor fit to residuals. I am aware of the time constraints placed on most assessment teams globally and that often bringing different institutions together (e.g. Government department and University) can be problematic due to different performance drivers and institutional expectations. Collaborations such as these should be encouraged and I congratulate the assessment team on pursuing the collaboration. In my experience, these collaborations normally produce better outcomes but do take longer.

## Appendix 1: Bibliography of materials provided for review

### References

- Cadrin SX, Boutillier JA, & Idoine JS. 2004. A hierarchical approach to determining reference points for Pandalid shrimp. *Canadian Journal of Fisheries and Aquatic Sciences* 61(8): 1373-1391. 19 p.
- Chen Y, Hunter M, Vadas R, Beal B. 2003. Developing a growth-transition matrix for the stock assessment of the green sea urchin (*Strongylocentrotus droebachiensis*) off Maine. *Fishery Bulletin* 101:737-744. 8 p.
- Chen Y, Kanaiwa M, & Wilson C. 2005. Developing and evaluating a size-structured stock assessment model for the American lobster, *Homarus americanus*, fishery. *New Zealand Journal of Marine and Freshwater Research* (Special issue on Lobster Biology and Management). 39:645-660. 16 p.
- Clark SH, Cadrin SX, Schick DF, Diodati PJ, Armstrong MP, and McCarron D. 2000. The Gulf of Maine Northern Shrimp (*Pandalus borealis*) Fishery: a Review of the Record. *J. Northw. Atl. Fish. Sci.* 27: 193-226. 34 p.
- Cotter J. 2009. Northeast Stock Assessment Review Committee (SARC 49) for Atlantic Surfclams in the US EEZ and Butterfish. 30 Nov – 3 Dec 2009. Woods Hole, MA. 49 p.
- Link JS & Idoine JS. 2009. Estimates of Predator Consumption of the Northern Shrimp *Pandalus borealis* with Implications for Estimates of Population Biomass in the Gulf of Maine. *North American Journal of Fisheries Management*, 29(6): 1567-1583. 18 p.
- McBride et al. 2013. Changes in size and age at maturity of the northern stock of Tilefish (*Lopholatilus chamaeleonticeps*) after a period of overfishing. *Fish Bull.* 111: 161-174. 14 p.
- Marsh et al. 1999. The 1882 tilefish kill - a cold event in shelf waters off the north-eastern United States? *Fish. Oceanogr.* 8:1, 39-49, 1999. 11 p.
- Northeast Fisheries Science Center. 2007. 45th Northeast Regional Stock Assessment Workshop (45th SAW): 45th SAW assessment report. US Dep Commer, Northeast Fish Sci Cent Ref Doc 07-16; 370 p.
- Northeast Fisheries Science Center. 2009. 48th Northeast Regional Stock Assessment Workshop (48th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 09-15; 834 p.
- Northeast Fisheries Science Center. 2009. 48th Northeast Regional Stock Assessment Workshop (48th SAW) assessment summary report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 09-10; 50 p.
- Northeast Fisheries Science Center. 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Summary Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-01; 41 p.

- Northeast Fisheries Science Center. 2010. 49th Northeast Regional Stock Assessment Workshop (49th SAW) Assessment Report. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 10-03; 383 p.
- Northeast Fisheries Science Center. 2009. Summary Report of the 49th Northeast Regional Stock Assessment Review Committee (SARC 49). 30 Nov – 3 Dec 2009. Woods Hole, MA. 37 p.
- Northeast Fisheries Science Center. 2013. Butterfish model diagnostics for the ASAP3 base model which served as the basis for developing the final model for SAW58 in 2014. 73p.
- Richards RA, Fogarty MJ, Mountain DG, Taylor MH. 2012. Climate change and northern shrimp recruitment variability in the Gulf of Maine. Marine Ecology Progress Series 464:167-178. 12 p.
- Smith MT. Northeast Regional Stock Assessment Review Committee (SARC 49). 2009. Independent reviewer's report on the 2009 Atlantic Surfclam and Butterfish Benchmark Stock Assessments. December 2009. 30 Nov – 3 Dec 2009. Woods Hole, MA. 51 p.
- Sparholt H. Northeast Regional Stock Assessment Review Committee (SARC 49). 2009. SARC 49: Butterfish and Atlantic surfclam Benchmark Stock Assessments. 30 Nov – 3 Dec 2009. Woods Hole, MA. 33 p.

### **Working Papers**

- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Report of Butterfish. Working Paper #1. SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.
- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Report of Tilefish. Working Paper #1.SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.
- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Report of Northern shrimp. Working Paper #1.SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.
- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Summary Report of Butterfish. Working Paper #2. SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.
- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Summary Report of Tilefish. Working Paper #2.SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.
- Working Group, Stock Assessment Workshop (SAW 58) 2014. Stock Assessment Summary Report of N. shrimp. Working Paper #2.SAW/SARC 58. January 27-31, 2014, NOAA Fisheries, Northeast Fisheries Science Center. Woods Hole, MA.

## Appendix 2: A copy of this Statement of Work

### Statement of Work

**58th Stock Assessment Workshop/Stock Assessment Review Committee (SAW/SARC): Benchmark stock assessments for butterfish, tilefish, and northern shrimp**

***Statement of Work (SOW) for CIE Panelists***

***(including a description of SARC Chairman's duties)***

### BACKGROUND

The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are independently selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (TORs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from [www.ciereviews.org](http://www.ciereviews.org).

### SCOPE

**Project Description:** The Northeast Regional Stock Assessment Review Committee (SARC) meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The SARC is the cornerstone of the Northeast Stock Assessment Workshop (SAW) process, which includes assessment development (SAW Working Groups or ASMFC technical committees), assessment peer review, public presentations, and document publication. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results provide the scientific basis for fishery management in the northeast region.

#### **Brief description of the science to be peer reviewed, and its relevant importance:**

The purpose of this meeting will be to provide an external peer review of benchmark stock assessments for **butterfish, tilefish, and northern shrimp**. This review determines whether the scientific assessments are adequate to serve as a basis for developing fishery management advice. Results form the scientific basis for fishery management in the northeast region.

## OBJECTIVES

The SARC review panel will be composed of three appointed reviewers from the Center of Independent Experts (CIE), and an independent chair from the SSC of the New England or Mid-Atlantic Fishery Management Council. The SARC panel will write the SARC Summary Report and each CIE reviewer will write an individual independent review report.

Duties of reviewers are explained below in the “**Requirements for CIE Reviewers**”, in the “**Charge to the SARC Panel**” and in the “**Statement of Tasks**”. The stock assessment Terms of Reference (TORs) are attached in **Annex 2**. The draft agenda of the panel review meeting is attached in **Annex 3**. The SARC Summary Report format is described in **Annex 4**.

**Requirements for the reviewers:** Three reviewers shall conduct an impartial and independent peer review of the striped bass and summer flounder stock assessments, and this review should be in accordance with this SoW and stock assessment TORs herein. The reviewers shall have working knowledge and recent experience in the application of modern fishery stock assessment models. Expertise should include statistical catch-at-age, state-space and index methods. Reviewers should also have experience in evaluating measures of model fit, identification, uncertainty, and forecasting. Reviewers should have experience in development of Biological Reference Points that includes an appreciation for the varying quality and quantity of data available to support estimation of Biological Reference Points. SARC 58 will address fishery stock assessments of **butterfish, tilefish, and northern shrimp**. For shrimp and butterfish, experience in the following is desirable: assessment of short-lived species, stocks where the environment and environmental change can impact recruitment and availability in research surveys. Specifically for tilefish: experience with assessments based on commercial catch per unit of effort.

## PERIOD OF PERFORMANCE

The contractor shall complete the tasks and deliverables as specified in the schedule of milestones within this statement of work. Each reviewer’s duties shall not exceed a maximum of 16 days to complete all work tasks of the peer review described herein.

Not covered by the CIE, the SARC chair’s duties should not exceed a maximum of 16 days (i.e., several days prior to the meeting for document review; the SARC meeting in Woods Hole; several days following the open meeting for SARC Summary Report preparation).

## PLACE OF PERFORMANCE AND TRAVEL

Each reviewer shall conduct an independent peer review during the panel review meeting scheduled in Woods Hole, Massachusetts during dates of January 27-31, 2014.

## STATEMENT OF TASKS

**Charge to SARC panel:** During the SARC meeting, the panel is to determine and write down whether each stock assessment Term of Reference (TOR) of the SAW (see **Annex 2**) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried

out correctly, and the conclusions are correct/reasonable. **If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted.** Where possible, the SARC chair shall identify or facilitate agreement among the reviewers for each stock assessment Term of Reference of the SAW.

If the panel rejects any of the current BRP or BRP proxies (for  $B_{MSY}$  and  $F_{MSY}$  and  $MSY$ ), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.

Each reviewer shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

**Tasks prior to the meeting:** The contractor shall independently select qualified reviewers that do not have conflicts of interest to conduct an independent scientific peer review in accordance with the tasks and TORs within the SoW. Upon completion of the independent reviewer selection by the contractor's technical team, the contractor shall provide the reviewer information (full name, title, affiliation, country, address, email, phone number, FAX number, and a CV suitable for the public) to the COR, who will forward this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The contractor shall be responsible for providing the SoW and stock assessment TORs to each reviewer. The NMFS Project Contact will be responsible for providing the reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact will also be responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or TORs must be made through the COR prior to the commencement of the peer review.

Foreign National Security Clearance: The reviewers shall participate during a panel review meeting at a government facility, and the NMFS Project Contact will be responsible for obtaining the Foreign National Security Clearance approval for the reviewers who are non-US citizens. For this reason, the reviewers shall provide by FAX (or by email if necessary) the requested information (e.g., 1.name [first middle and last], 2.contact information [address, telephone number], 3.gender, 4.country of birth, 5.country of citizenship, 6.country of permanent residence, 7.whether there is dual citizenship, 8.country of current residence, 9.birth date [mo, day, year], 10.passport number, 11.country of passport) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: <http://deemedexports.noaa.gov/>.

Pre-review Background Documents and Working Papers: Approximately two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the SARC chair and CIE reviewers the necessary background information and reports (i.e., working papers prepared by the SAW Working Group) for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the COR on where to send documents. The reviewers are responsible only for the pre-review documents that are delivered to the

contractor in accordance to the SoW scheduled deadlines specified herein. The reviewers shall read all documents deemed as necessary in preparation for the peer review.

**Tasks during the panel review meeting:** Each reviewer shall conduct the independent peer review in accordance with the SoW and stock assessment TORs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and TORs shall not be made during the peer review, and any SoW or TORs modifications prior to the peer review shall be approved by the COR and contractor.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the stock assessment TORs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

(SARC chair)

Act as chairperson, where duties include control of the meeting, coordination of presentations and discussions, making sure all stock assessment Terms of Reference of the SAW are reviewed, control of document flow, and facilitation of discussion. For each assessment, review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to discuss the stock assessment and to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

(SARC CIE reviewers)

For each stock assessment, participate as a peer reviewer in panel discussions on assessment validity, results, recommendations, and conclusions. From a reviewer's point of view, determine whether each stock assessment Term of Reference of the SAW was completed successfully. Terms of Reference that are completed successfully are likely to serve as a basis for providing scientific advice to management. If a reviewer considers any existing Biological Reference Point or BRP proxy to be inappropriate, the reviewer should try to recommend an alternative, should one exist. Review both the Assessment Report and the draft Assessment Summary Report. The draft Assessment Summary Report is reviewed and edited to assure that it is consistent with the outcome of the peer review, particularly statements that address stock status and assessment uncertainty.

During the question and answer periods, provide appropriate feedback to the assessment scientists on the sufficiency of their analyses. It is permissible to request additional information if it is needed to clarify or correct an existing analysis and if the information can be produced rather quickly.

## Tasks after the panel review meeting:

### SARC CIE reviewers:

Each CIE reviewer shall prepare an Independent CIE Report (see Annex 1). This report should explain whether each stock assessment Term of Reference of the SAW was or was not completed successfully during the SARC meeting, using the criteria specified above in the “Charge to SARC panel” statement.

If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent CIE Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.

During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments may be raised. Comments on these questions should be included in a separate section at the end of the Independent CIE Report produced by each reviewer.

The Independent CIE Report can also be used to provide greater detail than the SARC Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

### SARC chair:

The SARC chair shall prepare a document summarizing the background of the work to be conducted as part of the SARC process and summarizing whether the process was adequate to complete the stock assessment Terms of Reference of the SAW. If appropriate, the chair will include suggestions on how to improve the process. This document will constitute the introduction to the SARC Summary Report (see Annex 4).

### SARC chair and CIE reviewers:

The SARC Chair, with the assistance from the CIE reviewers, will prepare the SARC Summary Report. Each CIE reviewer and the chair will discuss whether they hold similar views on each stock assessment Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the SAW. For terms where a similar view can be reached, the SARC Summary Report will contain a summary of such opinions. In cases where multiple and/or differing views exist on a given Term of Reference, the SARC Summary Report will note that there is no agreement and will specify - in a summary manner – what the different opinions are and the reason(s) for the difference in opinions.

The chair’s objective during this SARC Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. The chair will take the lead in editing and completing this report. The chair may express the chair’s opinion on each Term of Reference of the SAW, either as part of the group opinion, or as a separate minority opinion.

The SARC Summary Report (please see Annex 4 for information on contents) should address whether each stock assessment Term of Reference of the SAW was completed successfully.



For each Term of Reference, this report should state why that Term of Reference was or was not completed successfully. The Report should also include recommendations that might improve future assessments.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, the SARC Summary Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRP proxies are the best available at this time.

The contents of the draft SARC Summary Report will be approved by the CIE reviewers by the end of the SARC Summary Report development process. The SARC chair will complete all final editorial and formatting changes prior to approval of the contents of the draft SARC Summary Report by the CIE reviewers. The SARC chair will then submit the approved SARC Summary Report to the NEFSC contact (i.e., SAW Chairman).

## DELIVERY

Each reviewer shall complete an independent peer review report in accordance with the SoW. Each reviewer shall complete the independent peer review according to required format and content as described in **Annex 1**. Each reviewer shall complete the independent peer review addressing each stock assessment TOR listed in **Annex 2**.

**Specific Tasks for CIE Reviewers:** The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

- 1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
- 2) Participate during the panel review meeting at the Woods Hole, Massachusetts scheduled during January 27-31, 2014.
- 3) Conduct an independent peer review in accordance with this SoW and the assessment TORs (listed in **Annex 2**).
- 4) No later than February 14, 2014, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shrivani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and to Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in **Annex 1**, and address each assessment TOR in **Annex 2**.

**Schedule of Milestones and Deliverables:** The contractor shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

|                   |   |
|-------------------|---|
| December 16, 2013 | Contractor sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact |
|-------------------|---|

|                     |   |
|---------------------|---|
| January 13, 2014    | NMFS Project Contact will attempt to provide reviewers the pre-review documents   |
| January 27-31, 2014 | Each reviewer participates and conducts an independent peer review during the panel review meeting in Woods Hole, MA    |
| January 31, 2014    | SARC Chair and CIE reviewers work at drafting reports during meeting at Woods Hole, MA, USA                             |
| February 14, 2014   | Reviewers submit draft independent peer review reports to the contractor's technical team for independent review        |
| February 17, 2014   | Draft of SARC Summary Report, reviewed by all CIE reviewers, due to the SARC Chair *                                    |
| February 21, 2014   | SARC Chair sends Final SARC Summary Report, approved by CIE reviewers, to NEFSC contact (i.e., SAW Chairman)            |
| February 28, 2014   | Contractor submits independent peer review reports to the COR who reviews for compliance with the contract requirements |
| March 7, 2014       | The COR distributes the final reports to the NMFS Project Contact and regional Center Director                          |

\* The SARC Summary Report will not be submitted, reviewed, or approved by the CIE.

The SAW Chairman will assist the SARC chair prior to, during, and after the meeting in ensuring that documents are distributed in a timely fashion.

NEFSC staff and the SAW Chairman will make the final SARC Summary Report available to the public. Staff and the SAW Chairman will also be responsible for production and publication of the collective Working Group papers, which will serve as a SAW Assessment Report.

**Modifications to the Statement of Work:** Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on substitutions. The COR can approve changes to the milestone dates, list of pre-review documents, and TORs within the SoW as long as the role and ability of the reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and TORs shall not be changed once the peer review has begun.

**Acceptance of Deliverables:** The deliverables shall be the final peer review report from each reviewer that satisfies the requirements and terms of reference of this SoW. The contract shall be successfully completed upon the acceptance of the contract deliverables by the COR based on three performance standards:

- (1) each report shall be completed with the format and content in accordance with **Annex 1**,
- (2) each report shall address each stock assessment TOR listed in **Annex 2**,

(3) each report shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Upon the acceptance of each independent peer review report by the COR, the reports will be distributed to the NMFS Project Contact and pertinent NMFS science director, at which time the reports will be made publicly available through the government's website.

The contractor shall send the final reports in PDF format to the COR, designated to be William Michaels, via email William.Michaels@noaa.gov

**Support Personnel:**

William Michaels, Program Manager, COR  
NMFS Office of Science and Technology  
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910  
William.Michaels@noaa.gov Phone: 301-427-8155

Manoj Shivilani, CIE Lead Coordinator  
Northern Taiga Ventures, Inc.  
10600 SW 131<sup>st</sup> Court, Miami, FL 33186  
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President  
Northern Taiga Ventures, Inc. (NTVI)  
22375 Broderick Drive, Suite 215, Sterling, VA 20166  
RPerretti@ntvifederal.com Phone: 571-223-7717

**Key Personnel:**

Dr. James Weinberg, NEFSC SAW Chairman, NMFS Project Contact  
Northeast Fisheries Science Center  
166 Water Street, Woods Hole, MA 02543  
James.Weinberg@noaa.gov (Phone: 508-495-2352) (FAX: 508-495-2230)

Dr. William Karp, NEFSC Science Director  
Northeast Fisheries Science Center  
166 Water St., Woods Hole, MA 02543  
william.karp@noaa.gov Phone: 508-495-2233

## **Annex 1: Format and Contents of Independent Peer Review Report**

1. The independent peer review report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).
2. The main body of the report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Findings of whether they accept or reject the work that they reviewed, and an explanation of their decisions (strengths, weaknesses of the analyses, etc.) for each TOR, and Conclusions and Recommendations in accordance with the TORs. For each assessment reviewed, the report should address whether each TOR of the SAW was completed successfully. For each TOR, the Independent Review Report should state why that TOR was or was not completed successfully. To make this determination, the SARC chair and reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice.
  - a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.
  - b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, and especially where there were divergent views.
  - c. Reviewers should elaborate on any points raised in the SARC Summary Report that they feel might require further clarification.
  - d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.
  - e. The independent report shall be a stand-alone document for others to understand the proceedings and findings of the meeting, regardless of whether or not others read the SARC Summary Report. The independent report shall be an independent peer review of each TOR, and shall not simply repeat the contents of the summary report.
3. The reviewer report shall include the following appendices:
  - Appendix 1: Bibliography of materials provided for review
  - Appendix 2: A copy of this Statement of Work
  - Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

## **Annex 2: 58<sup>th</sup> SAW/SARC Stock Assessment Terms of Reference**

(file vers.: 8/2/2013)

### **A. Butterfish**

1. Characterize the commercial catch including landings, effort and discards by gear type. Describe the magnitude of uncertainty in these sources of data.
2. Characterize the survey data that are being used in the assessment. Describe the magnitude of uncertainty in these sources of data.
3. Characterize oceanographic and habitat data as it pertains to butterfish distribution and availability. If possible, integrate the results into the stock assessment (TOR-5).
4. Evaluate consumptive removals of butterfish by its predators. If possible, integrate results into the stock assessment (TOR-5).
5. Use assessment models to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Include a comparison with previous assessment results and previous projections.
6. State the existing stock status definitions for “overfished” and “overfishing”. Given that the stock status is currently unknown, update or redefine biological reference points (BRPs; point estimates for  $B_{MSY}$ ,  $B_{THRESHOLD}$ ,  $F_{MSY}$  and  $MSY$ , or their proxies) and provide estimates of their uncertainty. Consider effects of environmental factors on stability of reference points and implications for stock status.
7. Evaluate stock status with respect to a newly proposed model and with respect to “new” BRPs and their estimates (from TOR-6). Evaluate whether the stock is rebuilt.
8. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).
  - a. Provide numerical annual projections (2 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for  $F$ , and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment). Comment on which projections seem most realistic.
  - b. Describe this stock’s vulnerability (see “Appendix to the SAW TORs”) to becoming overfished, and how this could affect the choice of ABC.
9. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

## B. Tilefish

1. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the magnitude of uncertainty in these sources of data.
2. Characterize commercial LPUE as a measure of relative abundance. Consider the utility of recreational data for this purpose. Characterize the uncertainty and any bias in these sources of data.
3. For the depth zone occupied by tilefish, examine the relationship between bottom temperature, tilefish distribution and thermal tolerance.
4. Use assessment models to estimate annual fishing mortality and stock size for the time series, and estimate their uncertainty. Include a historical retrospective to allow a comparison with previous assessment results.
5. State the existing stock status definitions for “overfished” and “overfishing”. Then update or redefine biological reference points (BRPs; point estimates for  $B_{MSY}$ ,  $B_{THRESHOLD}$ ,  $F_{MSY}$  and  $MSY$  or for their proxies) and provide estimates of their uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for BRPs. Comment on the scientific adequacy of existing BRPs and the “new” (i.e., updated, redefined, or alternative) BRPs.
6. Evaluate stock status with respect to the existing ASPIC model (from previous peer reviewed accepted assessment) and with respect to a new model developed for this peer review. In both cases, evaluate whether the stock is rebuilt.
  - a. When working with the existing model, update it with new data and evaluate stock status (overfished and overfishing) with respect to the existing BRP estimates.
  - b. Then use the newly proposed model and evaluate stock status with respect to “new” BRPs and their estimates (from TOR-4).
7. Develop approaches and apply them to conduct stock projections and to compute the statistical distribution (e.g., probability density function) of the OFL (overfishing level) and candidate ABCs (Acceptable Biological Catch; see Appendix to the SAW TORs).
  - a. Provide numerical annual projections (2-3 years). Each projection should estimate and report annual probabilities of exceeding threshold BRPs for  $F$ , and probabilities of falling below threshold BRPs for biomass. Use a sensitivity analysis approach in which a range of assumptions about the most important uncertainties in the assessment are considered (e.g., terminal year abundance, variability in recruitment).
  - b. Comment on which projections seem most realistic. Consider the major uncertainties in the assessment as well as sensitivity of the projections to various assumptions.
  - c. Describe this stock’s vulnerability (see “Appendix to the SAW TORs”) to becoming overfished, and how this could affect the choice of ABC.

8. Review, evaluate and report on the status of the SARC and Working Group research recommendations listed in most recent SARC reviewed assessment and review panel reports. Identify new research recommendations.

### **C. Northern shrimp**

1. Present the Gulf of Maine northern shrimp landings, discards, effort, and fishery-independent data used in the assessment. Characterize the precision and accuracy of the data and justify inclusion or elimination of data sources.
2. Estimate population parameters (fishing mortality, biomass, and abundance) using assessment models. Evaluate model performance and stability through sensitivity analyses and retrospective analysis, including alternative natural mortality ( $M$ ) scenarios. Include consideration of environmental effects where possible. Discuss the effects of data strengths and weaknesses on model results and performance.
3. Update or redefine biological reference points (BRPs; point estimates or proxies for  $B_{MSY}$ ,  $SSB_{MSY}$ ,  $F_{MSY}$ ,  $MSY$ ). Evaluate stock status based on BRPs.
4. Characterize uncertainty of model estimates of fishing mortality, biomass and recruitment, and biological reference points.
5. Review the methods used to calculate the annual target catch and characterize uncertainty of target catch estimates.
6. Develop detailed short and long-term prioritized lists of recommendations for future research, data collection, and assessment methodology. Highlight improvements to be made before the next benchmark assessment.
7. Based on the biology of species, and potential scientific advances, comment on the appropriate timing of the next benchmark assessment and intermediate updates.

## ***Appendix to the SAW Assessment TORs:***

### **Clarification of Terms used in the SAW/SARC Terms of Reference**

#### ***Appendix to the Assessment TORs:***

**Explanation of “Acceptable Biological Catch”** (DOC Natl. Standard Guidelines, Fed. Reg., vol. 74, no. 11, 1/16/2009):

*Acceptable biological catch (ABC)* is a level of a stock or stock complex’s annual catch that accounts for the scientific uncertainty in the estimate of [overfishing limit] OFL and any other scientific uncertainty...” (p. 3208) [In other words,  $OFL \geq ABC$ .]

*ABC for overfished stocks.* For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. (p. 3209)

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of “catch” that is “acceptable” given the “biological” characteristics of the stock or stock complex. As such, [optimal yield] OY does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

**Explanation of “Vulnerability”** (DOC Natl. Standard Guidelines, Fed. Reg., vol. 74, no. 11, 1/16/2009):

*“Vulnerability.* A stock’s vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce MSY and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality).” (p. 3205)

**Rules of Engagement among members of a SAW Assessment Working Group:**

Anyone participating in SAW assessment working group meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.



**Annex 3: Draft Agenda**  
**58th Northeast Regional Stock Assessment Workshop (SAW 58)**  
**Stock Assessment Review Committee (SARC) Meeting**  
**January 27-31, 2014**

Stephen H. Clark Conference Room – Northeast Fisheries Science Center  
Woods Hole, Massachusetts

**DRAFT AGENDA\* (version: 25 October 2013)**

| TOPIC                  | PRESENTER(S)                                  | SARC LEADER | RAPPORTEUR |
|------------------------|---|-------------|------------|
| <b>Monday, Jan. 27</b> |   |             |            |
| <b>10 – 10:30 AM</b>   |   |             |            |
| Welcome                | <b>James Weinberg</b> , SAW Chair             |             |            |
| Introduction           | <b>Robert Latour</b> , SARC Chair             |             |            |
| Agenda                 |   |             |            |
| Conduct of Meeting     |   |             |            |
| <b>10:30 – 12:30</b>   | Assessment Presentation (A. Butterfish)       |             |            |
|                        | <b>TBD</b>                                    | <b>TBD</b>  | <b>TBD</b> |
| <b>12:30 – 1:30 PM</b> | Lunch   |             |            |
| <b>1:30 – 3:00</b>     | Assessment Presentation (A. Butterfish)       |             |            |
|                        | <b>TBD</b>                                    | <b>TBD</b>  | <b>TBD</b> |
| <b>3:00 – 3:15</b>     | Break   |             |            |
| <b>3:15 – 5:15</b>     | SARC Discussion w/ Presenters (A. Butterfish) |             |            |
|                        | <b>Robert Latour</b> , SARC Chair             |             | <b>TBD</b> |
| <b>5:15 – 5:45</b>     | Public Comments (A. Butterfish)               |             |            |

| TOPIC | PRESENTER(S) | SARC LEADER | RAPPORTEUR |
|-------|--------------|-------------|------------|
|-------|--------------|-------------|------------|

---

## Tuesday, Jan. 28

|                        |   |             |            |            |
|------------------------|---|-------------|------------|------------|
| <b>8:45 – 11 AM</b>    | Assessment Presentation (B. Tilefish)             | <b>TBD</b>  | <b>TBD</b> | <b>TBD</b> |
| <b>11:00 - 11:15</b>   | Break   |             |            |            |
| <b>11:15 – 12:15</b>   | SARC Discussion w/presenters (B. Tilefish)        |             |            |            |
|                        | <b>Robert Latour, SARC Chair</b>                  |             |            | <b>TBD</b> |
| <b>12:15 – 12 :30</b>  | Public Comments (B. Tilefish)                     |             |            |            |
| <b>12:30 – 1:45 PM</b> | Lunch   |             |            |            |
| <b>1:45 – 4:15</b>     | Assessment Presentation (C. Northern shrimp)      | <b>TBD</b>  | <b>TBD</b> |            |
| <b>4:15 - 4:30</b>     | Break   |             |            |            |
| <b>4:30 – 5:45</b>     | SARC Discussion w/presenters (C. Northern shrimp) |             |            |            |
|                        | <b>Robert Latour, SARC Chair</b>                  |             |            | <b>TBD</b> |
| <b>5:45 – 6:00</b>     | Public Comments (C. Northern shrimp)              |             |            |            |
| <b>7:00</b>            | (Social Gathering )                               |             |            |            |
| TOPIC                  | PRESENTER(S)                                      | SARC LEADER | RAPPORTEUR |            |

---

## Wed. Jan. 29

|                        |   |  |  |            |
|------------------------|---|--|--|------------|
| <b>9:00 – 11:15 AM</b> | Revisit with presenters (A. Butterfish) |  |  |            |
|                        | <b>Robert Latour, SARC Chair</b>        |  |  | <b>TBD</b> |
| <b>11:15 – 11:30</b>   | Break                                   |  |  |            |
| <b>11:30 – 12:30</b>   | Revisit with presenters (B. Tilefish)   |  |  |            |
|                        | <b>Robert Latour, SARC Chair</b>        |  |  | <b>TBD</b> |
| <b>12:30 – 1:30 PM</b> | Lunch                                   |  |  |            |

|                    |  |            |
|--------------------|--|------------|
| <b>1:30 -2:30</b>  | (cont) Revisit with presenters (B. Tilefish)<br><b>Robert Latour, SARC Chair</b> | <b>TBD</b> |
| <b>2:30 – 2:45</b> | Break  |            |
| <b>2:45 – 5:15</b> | Revisit with presenters (C. Northern shrimp)<br><b>Robert Latour, SARC Chair</b> | <b>TBD</b> |

### **Thur. Jan. 30**

|                         |  |            |
|-------------------------|--|------------|
| <b>8:30 – 11:30</b>     | Review/edit Assessment Summary Report (A. Butterfish)<br><b>Robert Latour, SARC Chair</b>      | <b>TBD</b> |
| <b>11:30 – 12:30 PM</b> | Lunch  |            |
| <b>12:30 – 2:45</b>     | Review/edit Assessment Summary Report (B. Tilefish)<br><b>Robert Latour, SARC Chair</b>        | <b>TBD</b> |
| <b>2:45 – 3:00</b>      | Break  |            |
| <b>3:00 - 5:30</b>      | Review/edit Assessment Summary Report (C. Northern shrimp)<br><b>Robert Latour, SARC Chair</b> | <b>TBD</b> |

### **Friday, Jan. 31**

|                          |                                       |
|--------------------------|---------------------------------------|
| <b>9:00 AM – 5:00 PM</b> | SARC Report writing. (closed meeting) |
|--------------------------|---------------------------------------|

\*All times are approximate, and may be changed at the discretion of the SARC chair. The meeting is open to the public, except where noted.

*The NMFS Project contact will provide the final agenda before the meeting.  
Reviewers must attend the entire meeting.*

#### **Annex 4: Contents of SARC Summary Report**

1.

The main body of the report shall consist of an introduction prepared by the SARC chair that will include the background, a review of activities and comments on the appropriateness of the process in reaching the goals of the SARC. Following the introduction, for each assessment reviewed, the report should address whether each Term of Reference of the SAW Working Group was completed successfully. For each Term of Reference, the SARC Summary Report should state why that Term of Reference was or was not completed successfully.

To make this determination, the SARC chair and CIE reviewers should consider whether the work provides a scientifically credible basis for developing fishery management advice. Scientific criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If the CIE reviewers and SARC chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

The report may include recommendations on how to improve future assessments.

2.

If any existing Biological Reference Points (BRP) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3.

The report shall also include the bibliography of all materials provided during the SAW, and relevant papers cited in the SARC Summary Report, along with a copy of the CIE Statement of Work.

The report shall also include as a separate appendix the assessment Terms of Reference used for the SAW, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

**Appendix 3: Panel Membership or other pertinent information from the panel review meeting.**

Robert J. Latour, Chair  
Catherine Dichmont  
Stewart Frusher  
Ian Jonsen